Adam M. Smith. Establishing User Needs for an Online Video Retrieval System. A Master's paper for the M.S. in I.S. degree. April, 2001. 59 pages. Advisor: Barbara M. Wildemuth

Individuals now have the ability to effectively access information such as images, audio and video over the Internet. However, it has been only recently that researchers have been able to dedicate themselves to establishing techniques for coping with the retrieval of these multidimensional non-textual documents.

This work examines some of the current work in the field of video retrieval and attempts to point out the importance of user needs analysis in retrieving video online. This study, through nine end-user interviews, takes a first step in providing information that designers can use in developing video retrieval systems. The results clearly indicate that there is a large need for video in many fields, and that there are common themes in the user needs within particular disciplines. Participants in this study wanted to express their queries in words, but needed to view stills and video clips when reviewing the results of their searches.

Headings:

Information systems – Special subjects – Video recordings

Information systems – Special subjects – Audiovisual materials

End-user searching – Aims and objectives

Information retrieval

Video recordings -- Cataloging

ESTABLISHING USER NEEDS FOR AN ONLINE VIDEO RETRIEVAL SYSTEM

by
Adam M. Smith

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

April, 2001

Approved by:

_____

Advisor

Introduction

The past decade has seen a tremendous amount of activity in the development of online information systems. The proliferation of the Internet has given individuals the ability to access vast amounts of widely distributed information from a single computer. The amount of accessible information as well as the number of people gaining access to this information is increasing daily. According to the Pew Internet and American Life Project (2001), in September 2000, the number of Web pages was growing at a rate of 5 million per day and the number of American adults with Internet access grew from about 88 million to more than 104 million in the second half of 2000.

The biggest challenge designers of information systems face is how to organize this vast amount of information so that people can locate the information they need for a particular purpose. This organization of information for retrieval involves many issues on numerous levels: indexing and assigning descriptive metadata, database design, network design, interface design and application programming. All of these topics are aspects of the information system design process, and each requires special attention in order to design a successful information system.

In the information and library science community, research involving the improvement of the retrieval performance of information systems has primarily concentrated on textual

documents. In the past 10 years, due to advances in compression and digitization techniques, bandwidth and network capacity, memory, storage space, and processing speed, it has become less agonizing to transfer larger and more complex data across networks. Individuals now have the ability to effectively access information such as images, audio and video over the Internet.

However, while there have been centuries of scholarship devoted to developing strategies for the effective organization and retrieval of textual documents, it has been only recently that researchers have been able to dedicate themselves to establishing techniques for coping with the retrieval of these multidimensional non-textual documents.

The Open Video Project

The current work is one piece of a funded project within the School of Information and Library Science at the University of North Carolina at Chapel Hill called The Open Video Project (OVP). In the words of the OVP team, the system is to be "a shared digital video repository and test collection. Active as a public site at www.open-video.org for more than a year, the collection currently contains about 350 video segments, representing more than 22 hours of footage." (Geisler, Marchionini, Nelson, Spinks, and Yang, in press, p. 2) A brief history of the development of this project is as follows:

> "The Open Video Project was developed at the Interaction Design Laboratory (IDL) at the School of Information and Library Science at the University of North Carolina at Chapel Hill. Building on initial contributions from U.S. government agencies, such as the National Archives and NASA, and from Carnegie Mellon's Informedia project, the Open Video collection currently

> contains several hundred digitized video segments consisting of more than 15GB of footage (primarily in MPEG-1 format). The collection is hosted as one of the first channels of the Internet2 Distributed Storage Infrastructure project, which supports distributed collection hosting for research and education in the Internet2 community. To make the video more widely available and to solicit contributions and feedback from the research community, we have published previous papers (Geisler, 2000; Slaughter et al., 2000) and are currently collaborating with the Open Archives Initiative (OAI) and with the National Institute of Standards and Technology (NIST) Text Retrieval Conference (TREC) team." (Geisler et al., 2)

By February 2001 a Microsoft Access database of video metadata was available on a WindowsNT server, and there were plans to transfer it to a MySQL database on a Linux server.

However, there were clear weaknesses in the database design. For example, the entire database consisted of a single flat-file table. There was little attention paid to normalization – there were many repeated values and the database was not scalable. This was troublesome because while the archive was small, new video submissions were being received periodically and it was expected that their archive would expand considerably. This original database design had served its purpose of making the video archive available as a test bed for user interface research – mainly to test video abstracting techniques. The database, initially, was of secondary importance to the team's primary research interest, which was to study how to present condensed video surrogate material most effectively to users.

Video abstraction is a very important aspect of video retrieval. People do not want to search through hundreds of hours of video attempting to determine what is a relevant

video and what is not. They want to be able to decide quickly whether a particular video will be useful. But how do you display what a video is “about” quickly and accurately? This is a huge problem related to retrieving video and one on which the developers of the Open Video Project wanted to concentrate initially.

Yet, in focusing on developing video abstracts to the exclusion of defining metadata and the querying process, a large part of the interface design was left to the team’s assumptions. Specifically, it was assumed that the user of the system has somehow been able to enter a query effectively using methods and metadata appropriate to his or her need and that a retrieval set of relevant video material has been returned. How will the user be able to enter a query? What choices will the user have with regard to search criteria? Only after a user has been able to formulate their query effectively via a suitable initial interface will the individual need to peruse video abstracts. If a user wishes to find video about English Bulldogs and only retrieves video of the English Parliamentary System, he or she will surely not need to look at video abstracts to discover they have not found what they were looking for.

So the goal of the work reported here is to establish what different kinds of video needs diverse users of video systems have. Why are people coming to retrieve video in the first place? What kinds of videos do they need and why? What sorts of metadata do they find useful in specifying the kinds of videos they need? What kinds of metadata will they find useful in determining which videos in a retrieval set are worth pursuing? Only after these questions are answered can a video abstracting technique become helpful.

User-Centered Design

Focusing on the needs of users is not a new avenue of research. Thirty years ago, in 1971, Wilfred J. Hansen, one of the first proponents of user-centered design, published a paper entitled, “User engineering principles for interactive systems.” In this paper he strongly expressed the need for interactive systems to be designed keeping the needs of the user in mind as well as making user effort as minimal as possible. He calls this user centered design style ‘user engineering,’ and in the paper describes his particular engineering principles. These principles have been summarized in Table 1.

Table 1
Hansen’s User Engineering Principles
The First Principle: Know the User

Main and Sub Principles	Description
1. Minimize Memorization	
• Selection Not Entry	Rather than type a character string or operation name, the user should select the appropriate item from a list displayed by the computer.
• Names Not Numbers	When the user is to select from a set of items he should be able to select among them by name. This avoids having to remember and special codes.
• Predictable Behavior	The user should be able to gain an ‘impression’ of the system and then understand its behavior in terms of that impression.

(table continues)

Main and Sub Principles	Description
• Access to System Information	The user should be given access to the various controlling parameters and should be able to modify from the console any parameter that he can modify in any other way.
2. Optimize Operations	
• Rapid Execution of Common Operations	Data structures should be chosen that optimize frequent operations.
• Display Inertia	The display should change as little as necessary to carry out a user's request.
• Muscle Memory	The system should be designed in such a way that very repetitive operations should be delegated not by the conscious mind, but by the lower part of the brain.
• Reorganize Command Parameters	Frequent user commands should be as convenient as possible, while infrequent commands can be relegated to subcommands.
3. Engineer for Errors	
• Good Error Messages	Error messages should be as specific as possible and not waste a user's time.
• Engineer out the Common Errors	If an error occurs frequently, it is not the fault of the user, it is a problem in the system design.
• Redundancy	The system should provide more than one means to an end.
• Data Structure Integrity	Regardless of system or hardware trouble some version of the user information will always be available.

Hansen admits that other disciplines had been working in this area previously, but with a slightly different perspective.

> "Disciplines similar to user engineering have been called human engineering, human factors, and ergonomics, but these terms most often refer to analog systems like airplane cockpits where the pilot guides a process. User

> engineering applies to digital systems where the goal is to store or retrieve information" (Hansen, 1971, p. 523).

Hansen was not the only scholar concerned with this avenue of research. He also points to several articles written around the same time by scholars with a similar focus on the user. "D. Engelbart refers to these principles as 'User Feature Design.' His point is that this term emphasizes that the features are being designed for the user rather than the other way around"(Hansen, 1971, p. 523). Hansen goes on to note other important works that discuss user-centered design, specifically that "other sets of user engineering principles have been reported by L.B. Smith and J.G. Mitchell…[and that] the reader should also read R.B. Miller's paper" (Hansen, 1971, p. 523).

In the 1980's, as computer usage became more pervasive, user-centered design became a greater topic of interest among researchers. However, as Brenda Dervin and Michael Nilan pointed out (1986, p. 150), there was "a major tension between information science research and practice." This tension, they felt, "results from the charge that studies have not informed practice" (1986, p. 150). In reviewing post-1978 Information Science literature, they observed that while there had been a substantial number of articles written pressing "to make information needs and uses a central focus of information systems" (1986, p. 151), there had not been a similar push among individuals actually designing the information systems. They illustrated this conceptual push toward user-centered design with two quotes from contemporary articles of the time:

> "… it becomes increasingly clear that the success of information services is more likely to be achieved through adjusting the services to meet the specific

> needs of an individual rather than trying to adapt the individual user to match the wholesale output of an information system" (Garvey, Tomita, & Woolf, 1979, p. 256).
>
> "Effective transition into the information age will require switching from information systems that are technology and content driven to information systems that are user driven" (Mick, Lindsey, & Callahan, 1980, p. 355).

Dervin and Nilan mention several authors concerned with altering system orientations, including, Belkin (1984), Belkin et al. (1982a; 1982b), Breton, Cronin, Dervin (1981; 1981; 1983a; 1983b), Durrance (1984), Ford, Garvey et al., Jarvelin & Repo, Krikelas, Lowry, MacMullin & Taylor, Maron, Ofori-Dwumfuo, Paisley (1980), Robertson, Vermuelen, White, Williamson, and Wooster.

Norman and Draper (1986) further stressed the importance of designing with the focus on the user. In their introduction, they state that their collection "is a book about the design of computers, but from the user's point of view: *User Centered System Design*. The emphasis is on people, rather than technology, although the powers and limits of contemporary machines are considered in order to know how to take that next step from today's limited machines toward more user-centered ones" (1986, p. 2). They go on to explain how their book fits into the larger field of User Centered Design at the time:

> "This book is primarily an expression of a pluralistic approach, but if it has a common theme – a unity in its diversity – it is that human-computer interface design is not one small aspect of the main business of software design, nor will it be illuminated (let alone "solved") by a single methodology or technical innovation. To begin with, we do not wish to ask how to improve upon an interface to a program whose function and even implementation has already been decided. We wish to attempt User Centered System Design, to ask what the goals and needs of the users are, what tools they need, what kind of tasks

> they wish to perform, and what methods they would prefer to use. We would like to start with the users, and to work from there.
>
> Granted the premise of User Centered System Design, though, what follows? The more we study it, the bigger the subject seems to become. Pluralism is the result of the piecemeal recognition of more and more important aspects to the subject. We are at the point (in the mid 1980s) of realizing just how much bigger the problem is than has usually been acknowledged, but we are not within sight of a grand synthesis or a unifying theory. This book offers 'perspectives' – pluralistic voices laying claim to your attention. The authors contributing to this book interacted to a considerable extent during its writing. As a result, many mutual connections have been found and are mentioned in the chapters, but nothing like a single synthesis has yet been constructed. The main message remains that pluralism is necessary and appropriate at this stage of the field. The chapters reflect this pluralism implicitly, not by design. (1986, pp. 2-3)

It is clear that Norman and Draper, like Dervin and Nilan, are trying to reinforce the idea of designing from a user's perspective.

Shneiderman's (1988) text on interface design marks the next milestone in user-centered design. It is one of the first books to begin to move away from theory and move more towards a practical guide for user centered interface design. In recent years, the importance of this work has been further recognized through awards such as the ACM Special Interest Group on Documentation (SIGDOC).

In the preface to the book, "Fighting for the User", Shneiderman writes that by 1988,

> "Researchers have shown that redesign of the human-computer interface can make a substantial difference in learning time, performance speed, error rates, and user satisfaction. Information and computer scientists have been testing design alternatives for their impact on these human performance measures. Commercial designers recognize that systems that are easier to use will have a competitive edge in information retrieval, office automation, and personal computing.

> Programmers and quality assurance teams are becoming more cautious and paying greater attention to the implementation issues that guarantee high quality user interfaces. Computer center managers are realizing that they must play an active role in ensuring that the software and hardware facilities provide high quality service to their users.
>
> In short, the divers use of computers in homes, offices, factories, hospitals, electric power control centers, hotels, banks, and so on is stimulating widespread interest in human factors issues. Human engineering, which was seen as the paint put on at the end of a project, is now understood to be the steel frame on which the structure is built." (Shneiderman, 1988, p. v)

One important difference between this work and others written at the time is that this book not only reiterates the importance of designing from a user centered perspective, it brings together previous research in cognitive psychology, human factors, information science, computer science and cites specific studies that provide metrics about the value of user centered design.

Shneiderman stated that his goals in writing this book were to "encourage greater attention to the user interface and to help develop a more rigorous science of the user interface design. Designing the User Interface presents design issues, offers experimental evidence where available, and makes reasonable recommendations where suitable." (Shneiderman, 1988, p. vii)

By the 1990s, user-centered design had become established as an integral part of the field of human computer interaction (HCI). Emphasis was placed less on convincing people of its value and more on developing methodologies (like those introduced by Shneiderman). The field of human computer interaction had continued to develop and grow and there

were numerous articles and books being written about user-centered design -- including HCI textbooks and practical guidebooks.

At this point, user-centered design has become an inseparable part of information system design. Designers are no longer debating the value of user oriented design – there are now established techniques and guidelines which designers feel are crucial to follow in order to insure maximum usability.

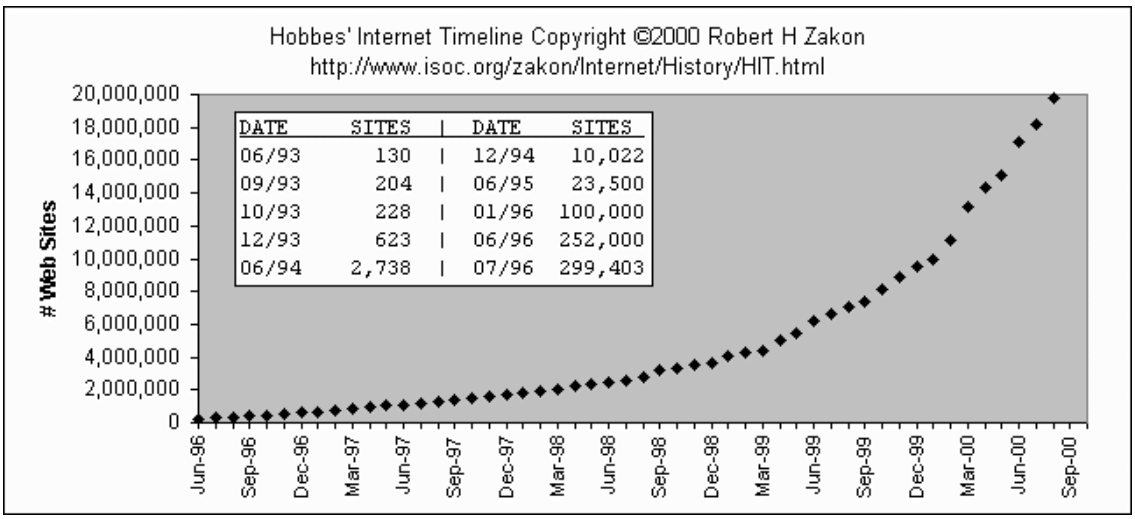
This concept of ‘usability’ is an important concept that arose out of the interest in user-centered design. The term implies and refers to a sense of measurement and shows that designers were attempting to quantify the effectiveness of their user-centered design techniques. Usability Engineering by Jakob Nielson,is to many an authoritative work on the practice of usability methodology. In the words of Amazon.com, “Jakob Nielsen's *Usability Engineering* provides a landmark guide to software design that has helped bring this area of research into the mainstream of computing” (Dragon). Nielsen writes in a very straightforward manner and his books on usability have quickly become standard texts in the field of usability engineering. Nielsen (1998), writing about his book, states that, “The basic philosophy of the book is YOU CAN DO IT! It is about cheap and fast methods that anybody can use in any interface design project (whether Web design, software design, or gadget design) to drastically improve usability. It is quite common to be able to cut users' learning time in half (thus cutting your training budget or support center costs by a similar amount).”

The concept of user-centered design began at least thirty years ago as a relatively novel concept. Scientists and researchers were designing systems that functioned but were not explicitly designed with the sanity of the user (particularly the non-technical user) in mind. As computers became more pervasive in our society, and as more and more people with different levels of computer experience and expertise began using these systems, the importance and value of systems that could be used easily and effectively became increasingly clear.

Scholars began to think about such questions as, “What makes a system usable?”, “How do people actually use systems?”, “How do we test to see what usability problems exist with our system?” They began to design practical techniques and methodologies that could be used to design user-centered systems. Moreover, they created tests that could be used to quantify the usability of the design of a system.

Designing for the World Wide Web

In the mid 1990s, an explosive event occurred which would have a dramatic effect on the field of human computer interaction and user-centered design – the creation and widespread use of the World Wide Web. The number of Web sites began to grow at an exponential rate. Figure 1 illustrates this point.



Note. From http://www.zakon.org/robert/internet/timeline
Figure 1. Hobbes' Internet Timeline, Copyright 2000 Robert H. Zakon. Reprinted with permission.

There is a great similarity between what happened with the advent of desktop publishing and the creation of Web site publishing systems. Individuals who were not trained as information designers (graphic designers in the desktop publishing example) were now creating their own information systems – Web sites. As more and more Web sites were being built and placed on the Internet, many people became exposed to the concept (although many didn't refer to it explicitly) of 'usability'. User-centered design was not just a concept to be considered by software and information system designers anymore. Now millions of individuals needed and wanted guidelines to follow in order to design the most usable Web sites they could.

Today, there has been explosive growth in the number of books devoted to Web site design and many specifically about usability. Some of the more popular include Jakob Nielsen (2000), Rosenfeld and Morville (1998), Jared Spool (1999), Deborah Mayhew

(1999), Hackos and Redish (1998), Bias and Mayhew (1994), and Mark Pearrow (2000). These represent only a few of the countless Web Design books currently on the market.

The World Wide Web, some scholars might say, has set back developments in HCI and interface design years because it has focused so many people's energies and forced them to design within the constraints of the World Wide Web. However, the Web has created a never before seen uproar over usability and user-centered design. In the long run, as more work is done in the field of user-centered design, and as we learn more about how to better measure usability and develop metrics to measure the benefits of user-centered design, software and information systems design will only benefit.

User Needs and Video Retrieval

Given that user-centered design has been a concern in the information system development community for the past thirty years, it was a little surprising to find that there were no extensive user analysis studies pertaining to video retrieval systems. While there have been many user studies which focus on testing and gathering data on particular interface design aspects of video retrieval, there is a scarcity of research dedicated to general user needs assessment regarding how and why people want to use online video systems.

The developers of the Open Video Project, as mentioned earlier, were mainly concerned with how users react to different video abstracting techniques. This browsing activity

takes place only after a query has been entered and information (including video and non-video information) has been retrieved and displayed.

> "For the Open Video Project, we aim to create an interface that provides users with specific features and mechanisms that are optimized to more easily browse and retrieve video. We believe it is especially crucial to provide users with maximum information to inform relevance judgments before accepting the time costs of downloading video. Thus, in addition to the retrieval task, we aim to help people understand a video collection's structure, what is and is not available, and what attributes might be useful for retrieval purposes. We are also providing people with a range of surrogates and integrating these surrogates into an effective and efficient interface. Although some projects have integrated multiple surrogates (for example, Informedia (Wactlar, et. al, 1996) and CueVideo (Ponceleon, et. al, 1999)), we aim to create an environment that provides multiple surrogates at the collection level as well as at the item level.
>
> Exactly how we create surrogates and integrate them into the interface is largely dependent on the needs and characteristics of our user audience. What features of video will best help our users find potentially relevant video segments, and which features of those segments will best help them evaluate whether a given segment is appropriate enough for their purpose to download? Determining this requires an understanding of how humans process video information, the needs of our user audience, and the possibilities for indexing and abstracting video." (Geisler et al., p. 8)

This scholarship is very important, but as a user study it seems somewhat narrow in scope. While the OVP team is indeed concerned with designing the best system for the user, they really do not have a good idea about exactly who their 'user' is, or what kind of user would find their different abstracting techniques useful. It could be that one kind of user fitting one profile would find a particular technique useful, while another user would not.

In general, video retrieval research has ignored general needs analysis studies, and instead has concerned itself with very specific aspects of video retrieval. Jain and Hampapur (1994) note that “research in the area of digital video analysis and video databases has been focused on the fine grain manipulation of video” (p. 2)

Similarly, Jain (1997) again points out that, with regard to the development of video retrieval systems, “researchers have generally focused on issues related to their own discipline” (p. 30) This domain-centric outlook could hurt research. “Such a narrow view usually results in defining artificial problems and developing solutions that are of no practical application. It is essential that we define problems and then develop solutions ignoring existing boundaries of traditional disciplines.” (p. 30) In order to balance the vast amount of domain specific research being done in this field, and to avoid, as Jain suggests, the creation of artificial problems and impractical solutions, more needs analysis and user profile research needs to be done.

It would be impossible to cite all the relevant and interesting research currently being done in the field of video retrieval. However, the literature seems to mainly focus on four topics:

- Non-textual querying of video
- Automatic indexing and automatic content analysis
- Video abstraction and automatic summarization
- Networking and compression – transmission of video

It must be noted that many of these studies do have the underlying goal of helping the user, and of creating the most intuitive and meaningful system. Yet they base their work on many assumptions about the users that may or may not be true. More often than not, one will see sentences that begin, “The user needs to be able to…,” or “The user needs tools that…,” or “This application satisfies the users’ needs by…” Unfortunately, there is no evidence regarding who is a member of the user population, what their characteristics are, or what information needs they bring to the retrieval process..

The following briefly summarizes some current work in the field of video retrieval:

Non-textual queries

There has been a good deal of work in the area of querying a system without the use of text. This has evolved due in large part to the extreme difficulty in developing an effective text-based search mechanism for images and video. Indexing video based on semantic content effectively and efficiently is one of the biggest challenges in the field today. “While there has been substantial progress with the presence of such systems as QBIC (Flicker, 1995), PhotoBook (Pentland, 1996), Virage (Hamrapur) and VisualSEEK (Smith 1996) most systems only support retrieval of still images.” (Chang, Chen, Meng, Sundaram, and Zhong, 1997)

Research in this area has developed systems where the user can query by example (QBE). With this method the user begins a query with a member of the database and uses it to search the database for other entities with similar attributes. Also, systems have been

developed where the user is allowed to sketch their query and then submit the drawing to the system to find similar images. There has also been work on allowing the user to specify motion to the system. Motion, time and audio combined are what make video such a unique and difficult medium to work with. It is natural, therefore, to try to allow users to attempt to specify these unique attributes. However, whether these methods and systems are particularly useful from a user's standpoint is yet to be seen. "Systems that can find sunset images using color histograms don't appreciably help a user who is really looking for a picture of Tom Cruise at the Oscars" (Chang, 1999, p. 313).

Automatic indexing and automatic content analysis

This field may be the most challenging and perhaps the most important to video retrieval. "Effective use of video in various applications is impeded by the difficulty of cataloging and managing video data. For example, media industries indicate that an hour long footage of video in the field can take up to tens hours to be fully cataloged and archived into the system" (Ponceleon, Srinivasan, Amir, Petkovic, & Diklic, 1998). Due to this prohibitive time cost, automatic indexing becomes crucial.

However, currently there is no way for a computer to understand what an image is "about." This remains solely a human capability. "The challenge, therefore, is that of designing a system that uses a balanced combination of automated techniques and human effort to populate, segment, and index the content of a digital library, enabling innovative search and browse interfaces to retrieve video collections" (Ponceleon et al., 1998).

Indexing video based on semantic content or 'aboutness' is an extremely challenging prospect. In reflecting upon the difficulty of indexing video and the current state of video retrieval systems Chang (1999) writes the following:

> "Only when we have automatic systems capable of video and audio understanding, not just similarity matching, will we be able to access multimedia by semantic content. Unfortunately, this problem seems to be impracticably difficult, if not impossibly AI-complete, and the most successful systems today use human-generated metadata such as closed captions."
>
> "The next stage in the evolution of multimedia information processing systems involves the extraction of semantic information. While intelligent algorithms will continue to increase the ability of machines to extract and understand information, these abilities will continue to lag far behind that of humans, especially for the understanding of non-linguistic information. Well designed human-machine interfaces that combine the intelligence of humans with the speed and power of computers will play a major role in creating a practical compromise between fully manual and completely automatic information systems" (p. 313).

Video abstraction and automatic summarization

While a computer does not have the ability to describe in words what an image or a video is "about," there is work being done to develop systems that automatically pull out the parts of a video that are deemed to be important or representative, in other words, to create a video abstract. One research group defines a video abstract as "a sequence of moving images, extracted from a longer video, much shorter than the original, and preserving the essential message of the original' (Lienhart, Pfeiffer, & Effelsberg, 1997). This area of study was, in fact, one of the original primary research interests of the OVP team.

The systems that have been developed typically break a video down into its parts: scenes, shots and frames. Exactly how this is done, and which shots or frames are pulled out as representative depends on the system. Furthermore, once the video has been broken down into its parts, and then ranked according to its importance, the problem then becomes how to display the summarized version to the user. While there is not enough space here to delve into all of the different techniques of displaying or the different algorithms used to break the video down, some of the most promising are described in Christel, Smith, Taylor and Winkler (1998), Ding, Marchionini, and Tse (1997), and Uchihashi, Foote, Girgensohn and Boreczky (1999).

Transmission of video

One of the biggest obstacles to effectively using a video retrieval system, especially online, are the problems of video compression and bandwidth. People want the highest quality of video possible, but they do not want to wait a long time for the video to download. Once again, there is a large amount of scholarship dedicated to the efficient transmission and storage of video, such as Haskins (1993), Keeton and Katz (1993), Rangan, Vin, and Ramanathan (1992), and Tobagi and Pang (1993).

User Needs and Query Formulation

The dearth of studies related to the reasons people need video has already been mentioned. However, a study by Rowe, Boreczky and Eads (1994) is somewhat related and has some of the same goals in mind as this study. In trying to design and implement a metadata database and query interface for video retrieval, they also noticed the problem

of not understanding who their users were and not knowing the kind of queries that would be asked of their system. They state that their “first problem was to characterize the types of queries that users want to ask” (1994). Similarly, they also find that “it is impossible to design a system to answer queries without determining the types of questions users will ask” (1994).

The main difference between this study and the one reported by Rowe, is that they asked potential users to come up with queries for a specific database of course lectures as well as for a database of motion pictures. In this research, the contents of the video database were not specified to the participants. It was left to the individual users, in the current research, to imagine a database that would provide for their realistic needs. By starting with the need and not the data set, it was hoped that light would be shed not only on the queries, but also on what kinds of data would be useful.

Methodology

The purpose of this research is to establish basic user needs as they pertain to online video retrieval. In order to gather data on user needs, interviews were conducted with potential users of online video retrieval systems. Potential users include people that would realistically have a need for and use video online. In choosing participants to contact, the biographies, publications, and research interests of faculty of the University of North Carolina at Chapel Hill were researched and twenty-eight people were identified who expressed an interest in using video. Nine people responded in time to be interviewed for this paper – one from Communication Studies, five from the School of

Journalism and Mass Communication, and three from the Computer Science Department. While this is indeed a small convenience sample, the participants interviewed did provide a good sense of the variety of the types of video that people need, as well as examples of how video is currently being used. Knowing these factors is vital in order to develop a successful video retrieval system.

By interviewing actual people about their video needs, three main questions were addressed:

- Why do people need video (especially online)?
- What metadata will they find useful in searching for video?
- What metadata will they find useful in distinguishing potential relevance of video in a retrieval set?

In addition, information was gathered on potential retrieval set interface designs – particularly how users might like to see video abstracts displayed (if at all).

The following are the specific questions that were asked of the participants (see Appendix A for the complete interview schedule:

1. What was your most recent experience with searching for video? What kind of video did you need and why? Was it a positive experience? Did you find what you were looking for?
2. In a few sentences, describe a realistic situation where you would possibly need to search for a video or video segment online? What kind of video would you need and why?

3. How important would specifying the following be in searching for your video?

   Title, Creator, Subject or keywords, Contributor, Creation date, Format, Language, Rights Information, File Size, Length, Amount of motion, Color, Sound, People

   (Each field was rated on a four point scale, Very useful to No use at all.)

   Are there any other criteria you would find useful in specifying the kind of video you need? Using the same scale as above, how useful would it be?

   For the video you are searching for, how useful would it be to be able to search using a technique other than either typing or clicking on words? For example:

   Drawing your query, Specifying colors, Specifying motion

   Are there any other non-textual techniques you can imagine would be useful in searching for the video you need? Using the same scale as above, how useful would it be?

4. For the next set of questions, imagine that you have input some kind of search criteria into the system and have received back a set of videos and corresponding information. How important would knowing the following be in determining whether a video is right for you:

   Title, Creator, Subject or keywords, Contributor, Creation date, Format, Language, Rights Information, File Size, Length, Amount of motion, Color, Sound, People

   (Each field was rated on a four-point scale, Very useful to No use at all.)

5. Regarding the above, would you rather retrieve video segments (smaller pieces of a video) and corresponding information or the entire video and corresponding information? If so, how would you like the video broken up? Using what criteria?

6. Would you want to be able to see still images of the video at this point? If so, how do you think you would like the images displayed?

7. Would you want to be able to see video clips at this point? If so, how do you think you would like the video displayed?

8. Would you want to be able to see written transcripts of the video if they were available at this point?

9. Would you want to be able to separate the videos you think you are interested in from the videos you think you are not interested in at this point?

10. Would you like to be able to use specific videos as examples in order to “search for similar videos?”

Given that the number of people interviewed was so small, the results should be viewed very tentatively. Therefore, the frequencies of the actual categorical responses regarding the usefulness of the specific metadata, while interesting, should not be seen as particularly reliable. However, the results, especially the additional comments offered by the participants are enlightening. The category response frequencies are provided in the next section, along with some of the more interesting comments.

Results and Discussion

Reasons for searching for videos

One of the most important aspects of this study was to find out why people needed video. The first question asked was intended to establish the most recent situation in which the participants were in need of video. This question also served to get the participants in the proper mindset regarding video retrieval. The video retrieval activity here did not necessarily have to involve an online system. The responses are summarized as follows:

Table 2 The Participant's Most Recent Video Needs

Participant ID	Most Recent Video Need
#1	Looked for recent controversial commercials on www.adcritic.com to show to a Journalism class.
#2	After being prompted by the Journalism Librarian, this professor ordered a video involving Daimler Chrysler advertisements.
#3	Had searched both in the Non-Print section of the library as well as online for videos that imitate, copy or pilfer other videotapes in Post-Modernist films. For example, American films that steal from foreign films. This was for this individuals scholarly research
#4	Searched online for old ads from the 50s and 60s to show to a Journalism class. Does not remember where they were found.
#5	Looked online for a specific academic lecture that dealt with performance modeling of the Internet.
#6	Looked for a movie trailer online.
#7	Went to www.atomfilms.com to watch short animations and films. This was purely for entertainment value.
#8	Needed footage of the Civil Rights Movement to put in a student news broadcast.
#9	Needed footage for a public television series on air pollution. Because of copyright laws, only the public television stations archives were searched.

The reasons the participants had most recently searched for video varied considerably. The journalism professors mostly wanted video of commercials to show their classes. This could be because in their field, which deals more with the medium of video than most, Students must learn how to use the medium of video and need to learn from seeing. While the journalism professors did express a strong desire for a system where they could

search and retrieve video quickly and easily, most expressed doubts about the feasibility as well as recounting past trouble with trying to use digital video in the classroom. Participant #4 stated that, “in my field, one of the biggest continuing problems is that you want to show students both old commercials as well as current commercials, and it is really hard to get a hold of clips that you don’t tape yourself on TV…It would be wonderful to be able to go online somewhere and grab a clip.”

While there is a demand from the journalism professors for this capability, many of them were pessimistic about the use of this technology in the classroom. Participant #4 recounted a past experience, “The trouble I have had with this is I can play them on my own computer, but somehow it always seems not to work when I try to take it into the classroom. Either saving it to a Zip doesn’t work or there is a problem with the network…something, I mean I am missing something in that process. To me, there should be a really fast way to do that, or some better way that I could go to the site in the classroom…my whole thing is that it should not take any more time than traditional methods. It should take a lot less time, but I am willing to do the same amount of time.” Similarly, participant #2 voiced concern over using technology in the classroom, “It would be nice to be able to pick up a commercial to illustrate points I am making…if I could do it easily and quickly. I know the capability exists in the classroom, but either A, I don’t know how to use it, or B, it doesn’t work.” Interestingly, participant #2 proceeded to attempt to find and play a commercial on a laptop from within the office and failed. This person reported “extreme frustration” with the technology.

Another interesting issue was raised by two participants who voiced strong concerns about copyright. Participant #6 said that instead of using real video to test infrastructures, his group usually winds up generating data that resembles video. “What we’ll do is synthetically generate something that, in the network, looks like that movie, but I mean it’s all junk data. Basically, the big issue that we are trying to stay away from is copyright violations.” Also, participant #9, who was working on producing a series that would be aired on public television, stated that his team was restricted to using video from the public television station’s own archives because of copyright issues. “The problem with working on a broadcast product that is going to be on air is all the copyright laws. Even if there was great video on CBS, you can’t use it. For the actual maker of the video, having other video is of no use whatsoever. It is really only of use to someone doing research, I would imagine.”

It is also worth noting briefly that two participants stated that their most recent experiences with watching video online were for pure entertainment reasons. One watched a movie trailer online, while the other went to atomfilms.com to watch animations and short films.

The next question that was asked was, in a way, an extension of the first. The participants were asked to describe a realistic scenario where they would need to retrieve video online. They were allowed to use the same answer to the first question if it suited them. The responses were:

Table 3 The Participant's Realistic Video Retrieval Scenarios

Participant ID	Realistic Video Retrieval Scenario
#1	Needed examples of television commercials to show a journalism class.
#2	Needed examples of television commercials to show a journalism class.
#3	Two scenarios were given: A. Needed videos for historical film research. B. Needed stock footage for documentary filmmaking.
#4	Needed television commercials to show to a journalism class.
#5	Needed to find videotaped seminar and lecture material online.
#6	Needed to find video that would stress network infrastructure for research purposes.
#7	Needed to find videotaped seminar and lecture material online.
#8	Needed stock footage for student news broadcasts.
#9	Needed video material for background information and leads for public television broadcast.

An interesting thing to note here is the similarity of answers among the participants. Three professors from the journalism department wanted the ability to search for and show video online to their classes. Also, two professors from the computer science department wanted the ability to search for and view lecture material online. Two other professors voiced a need for stock footage (one for personal documentary work, one for a student news broadcast). While this is a small sample of individuals and therefore difficult and perhaps unwise to draw generalizations, if this similarity of results were

found to be similar on a larger scale, it might drive the development of specialized departmental video archives.

Another finding of note is that most of the participants stated that they would be coming to the system with a specific need. The journalism professors wanted to show their journalism class a specific commercial. Participant #1 said that he liked to provoke class discussions by sometimes showing controversial commercials to the class. For example, it was stated that, previously, they had "looked for the controversial Bush ad with the 'Rats' that was during the presidential campaign. We did that one and there was another controversial ad that Nike did that showed a woman running through the woods with some chainsaw guy running after her."

The computer science professors stated that they were looking for specific lectures. Moreover, participant #7 stated that, "one scenario that I think will come up a lot and has come up a lot, is searching for learning and educational material, in particular, stuff that you've missed. That happens to me a lot. For example, I know there is a talk going on at some university, and I've missed it or can't go there…or even if they're broadcasting it, I don't have time that day…so I try and see if it's online somewhere."

Only two participants had a general need that wasn't content specific. Participant #3 expressed the need for stock footage for documentary work. This participant might go to a system with a general idea of what was needed, but would need to browse the system.

Participant #6, as was stated previously, simply needed complex video that would stress an infrastructure. This person couldn't care less what the actual content of the video was.

Video Metadata

The following table shows the results given to question three, involving which metadata people thought would be useful in specifying their specific query (in question 2) to an online system.

Table 4 Importance of metadata to search.
Question: How important would specifying the following be in searching for your video?

	Very Useful	Somewhat Useful	Not Very Useful	No Use at All
Title: the name given to the video.	1, 3A, 4, 5, 7, 8	2	9	3B, 6
Creator: the person or organization that created the video.	2, 3A, 7	1, 3B, 5, 9	8	4, 6
Subject/Keyword: terms that describe the video	1, 2, 3A, 3B, 4, 5, 6, 7, 8, 9			
Contributor: the person or organization that contributed the video to the collection.	2, 7, 9	1	3A, 3B, 4, 5, 8	6
Creation Date: the date the video was created.	7, 9	1, 3A, 5, 8	4	2, 3B, 6
Format: the digitized format of the video (example: MPEG, Real Video, Quicktime, etc.)	3A, 3B, 6, 9	4	1, 5, 7, 8	2
Language: the language spoken in the video.	5, 9	1, 3A, 3B	4, 7, 8	2, 6
Rights Information: copyright information and any restrictions of usage.	3B, 6, 9	1, 3A,	4, 7, 8	2, 5
File Size: the size of the video file.	3A, 3B	4, 5, 9	1, 7, 8	2, 6
Length: the length of the video file.	3A, 3B	1, 4, 9	2, 7, 8	5, 6
Amount of Motion: the amount of motion contained within a video on a scale.	6		1, 8, 9	2, 3A, 3B, 4, 5, 7
Color: whether or not the video is in color.	3B	3A, 9	1, 8	2, 4, 5, 6, 7
Sound: whether or not the video has sound.	3B, 4, 6, 9	1, 2, 3A, 8		5, 7
People: whether or not the video contains people.	8	2	1, 3A, 3B, 9	4, 5, 6, 7

The first conclusions one draws when looking at these results is that they vary widely. Given that the participants need video for very different reasons, this should come as no surprise. Another important finding is that all participants felt that searching using subject terms or keywords would be very useful. This finding suggests that the work currently being done involving automatic indexing of video is extremely important. Moreover, this response may illustrate that people feel most comfortable with presenting the system with their own words and descriptions instead of having to conform to the system or guess the system's linguistic abilities. While automatic indexing is an extremely difficult problem to solve for developers, the current results indicate that users feel strongly that they want to be able to search using text. So it seems that this problem is one that is extremely important to solve.

As noted earlier, most of the use scenarios involved the retrieval of a specific video. Generally speaking, the participants that wanted to find specific video believe that searching using Title and Subject/Keywords would be the most useful. Furthermore, many of the participants, when asked what other search criteria they would like to use to search a video system to satisfy their need, came up with types of criteria that were simply domain specific subject terms or keywords. The journalism professors that were searching for commercials stated that it would be useful to be able to specify the names of companies, genres of commercials (for example, humorous), country of origin, etc. The computer science professors that were looking for online lectures or talks, wanted to be able to specify the names of speakers, subject matter of lecture, location of lecture, etc.

Participant #3, who stated that if stock footage needed to be found, the following would be extremely useful to specify: camera movement (tracking shot, zoom shot, tilt shot, stationary), sound (stereophonic, monophonic), film formats (letterbox or widescreen versus standard). The thing that should be noted here is that the participants wanted to specify their needs to the system using terminology that was used in their field.

The individuals that were searching for less specific material, while they still found Subject/Keywords to be valuable, believed that specifying criteria like Format, Rights Information and Sound would be more valuable than specifying things like Title, Creator, and Creation date.

Non-textual querying

The next question, addressed the usefulness of non-textual querying. The results are shown in Table 5.

Table 5 Usefulness of Non-textual Querying for Search
Question: For the video you are searching for, how useful would it be to be able to search using a technique other than either typing or clicking on words? For example:

	Very Useful	Somewhat Useful	Not Very Useful	No Use at All
Drawing your Query			3	1, 2, 4, 5, 6, 7, 8, 9
Specifying Colors			3, 4	1, 2, 5, 6, 7, 8, 9
Specifying Motion			3, 4	1, 2, 5, 6, 7, 8, 9

None of the participants thought that the specified non-textual queries were particularly useful for their needs. Most participants doubted that they could express their needs without using words. For example, when asked whether drawing a query would be useful, participant #1 stated, “I wouldn’t be able to draw it so a computer would understand it.” Moreover, participant #4 stated, “I can’t imagine what I would draw.”

This illustrates the users’ eagerness to use words as the primary means of expressing their needs. This is, of course, intuitive. Language is the primary way in which humans express themselves – why should it be any different with human-computer interaction? A couple of participants, when asked whether they could think of any other kinds of non-textual queries, brought up voice recognition – although their ideas about how useful it would be were very different. Participant #2 stated that he would not find it useful at all because “I could type it as fast as I could speak it.” However, participant #4 was joyful at the prospect of being able to speak rather than type queries to a system. “Speech recognition would be awesome,” it was stated. “In a perfect world, I envision sitting at my desk and saying, ‘OK, I am thinking of an ad I saw 10 years ago and it had a car with a little kid…I think it was Volkswagen’…the computer would sort through and say, ‘Is it this one?’, and I can say, ‘No but it kind of looked like that.’ Then the computer would sort through and find all the other matches based on the hints…where we would work as partners.” When I asked how this was different than text querying, the response was, “Time and intuitiveness. Right now, when you have to type things in, there are all these parameters…or you screw up and it doesn’t work and you spent 15 minutes typing in something and then you lose the data and you have to start all over.” It is not really clear

what other participants felt about speech queries since it was not a interview question. However, this participant's responses suggest that some users would find it more intuitive to speak rather than type queries to a system.

These results raise the question of who would find this type of querying useful. It is possible that there are specific domains where these methods are useful, but there needs to be more work done which tries to identify the practical functionality of these methods. Participant #7, a professor within the computer science department, and a person that was familiar with the current research involving non-textual queries stated his reservations regarding these methods. "In fact, I have read a lot about QBIC papers and these things. They make for great research, I read a lot about them, I have gone to conferences…seen them presented…but unless I am a graphic artist who needs a particular visual feel, it's really hard to see how it is really going to help me. I think they are actually having a hard time."

Reviewing the results of a search

The next question asked the participant to imagine that they had somehow expressed a query to the system and had received a retrieval set. They were then asked how useful they felt the same criteria would be in distinguishing the importance or relevance of the returned videos. Their responses are summarized in Table 6.

Table 6 Importance of Metadata in Reviewing Results of Search

Question: For the next set of questions, imagine that you have input some kind of search criteria into the system and have received back a set of videos and corresponding information. How important would knowing the following be in determining whether a video is right for you:

	Very Useful	Somewhat Useful	Not Very Useful	No Use at All
Title: the name given to the video.	1, 2, 3, 5, 7	2, 4	9	6, 8
Creator: the person or organization that created the video.	3, 7	1, 2, 5, 9		4, 6, 8
Subject/Keyword: terms that describe the video.	1, 2, 4, 5, 7, 9	3		6, 8
Contributor: the person or organization that contributed the video to the collection.	7, 9	1, 2, 4	3, 5	6, 8
Creation Date: the date the video was created.	1, 3, 7, 9	2, 4, 5		6, 8
Format: the digitized format of the video (example: MPEG, Real Video, Quicktime, etc.)	3, 5, 9	7	1	2, 4, 6, 8
Language: the language spoken in the video.	1, 2, 3, 5, 9	7	1	2, 4, 6, 8
Rights Information: copyright information and any restrictions of usage.	1, 3B, 9	3A	2, 5, 7	4, 6, 8
File Size: the size of the video file.	3, 5	4, 7, 9	1	2, 6, 8
Length: the length of the video file.	3, 5	2, 4, 7, 9	1	6, 8
Amount of Motion: the amount of motion contained within a video on a scale.	6		1, 3, 9	2, 4, 5, 7, 8
Color: whether or not the video is in color.	3B, 9	3A	1	2, 4, 5, 6, 7, 8
Sound: whether or not the video has sound.	3B, 9	2, 3A, 4	1	5, 6, 7, 8
People: whether or not the video contains people.			1, 3, 9	2, 4, 5, 6, 7, 8

The value that the participants placed upon the displayed retrieval metadata closely resembles the values placed on querying metadata. There were some interesting things to point out however. More participants found Title, Creator, Subject/Keywords and Creation Date very useful at this point, than they did in the querying step. Also, more participants felt that Amount of Motion, Color, Sound, and People were of little or no use at this point.

Furthermore, a couple of respondents felt that at this point, any metadata that was presented was relatively unimportant. Specifically, the individual that wanted footage to supplement a news broadcast felt that the only thing that would be truly valuable at this point would be seeing the quality of the video and whether it actually contained the specific subject matter. Furthermore, the participant that needed video to simply stress infrastructure, stated that once a retrieval set was produced, the video would simply be plugged into the network and performance metrics would be produced – they would simply take what they got.

Question 5 examined the level of granularity that the participants desired in the items retrieved. Did the users need or want to see an entire video, or were they seeking out smaller, more specific segments within a larger video? The results indicate that it depends on the user's specific need. When asked whether or not searching on an entire video or a segment would be more useful, participant #7 summarized the situation stating, "I think it would be one or the other. Either I am looking for something specific because I want to watch the whole thing or I am looking to get a question answered and I am looking for

the information." Related to this, participant #7 further described how the process might evolve. "I might want to browse, you know, boring parts…skip forward. But if I am looking to answer a question, then I would want a segment…which is why I said automatic summarization would be very useful, because it could do both for me very easily. It would allow me to determine whether or not I want to watch the whole thing at the normal speed. In the other case it might answer my question quickly…and then if it doesn't, but it seems like something in the video would, then I could browse further."

Depending on what they were looking for, some of the participants wanted segments while others wanted the whole thing. For example, participant #1, who was looking for television commercials, said the whole video would be needed. Similarly, participant #5, who was looking for an online lecture, said that the entire video was needed. However, participant #3 stated that if a class was being taught in the art of filmmaking, or if stock footage was being sought for a documentary film, then it would be better to be able to search for specific shots that used particular film techniques. Participant #8, who was looking for specific civil rights footage to be used in a student news broadcast, stated that, "We would rather see specific stuff…like if we were to call up 'George Wallace,' and say, 'I want to see, specifically, George Wallace in the school house door.' I know what I am looking for. Give me only what I need and no more."

Another interesting thing to note was that while two of the participants, #4 and #9, stated that they would like to search for and retrieve specific segments, they still wanted to retain the ability to expand the segment and see more if they wanted. Participant #4 stated

that, “If you could retrieve a part and then you could say…if you could easily get more of the segment, then that would be helpful. What would not be helpful would be to decide you wanted a little segment, then you decided you actually wanted the whole thing and you had to go all the way back and start over.”

Interface Design

Question 6 focused on some possible interface design issues. The participants were asked whether or not they would like to see representative still images of the video at this point. While a few of the participants responded with rather apathetic yes/no answers, some felt that this functionality would be an important feature and further explained their thoughts. Participant #7 explained that the importance of still images at this point is very content dependent. “If it is a seminar then chances are the still images are just going to be a guy sitting in front of a blackboard. One caveat though, if the presenter is using a lot of visuals, like PowerPoint slides, then seeing the PowerPoint slides would be perfect.” This participant goes on to say that, with lecture material, “I am really looking for the audio quite honestly, [because] the video is not that interesting.”

One participant went on to describe how the still images should be displayed. Participant #4 stated that, “I would like to see a whole screen [of thumbnails] and then click on whichever ones looked like they might work…and enlarge it before I made a decision to do anything with it. I would like to be able to look at a whole lot at once. I think that would be faster…rather than going through them one at a time.”

In question 7, the participants were asked if they would like to be able to see video at this point. All of the participants felt that this would be very useful. Three participants, #5, #8, and #9 were interested in seeing the quality of the video. Participant #5, who was looking for online lecture material, stated that, “The main thing I want to see is what is the quality of this thing. You know, I don’t want to spend the cost of retrieving it and get back something jerky and of bad quality.” Participants #8 and #9, who were both involved in video production, felt that knowing the quality of the video was of paramount importance. Participant #8 wanted to see whether, “the video is well shot. Does it move all over the place? I can’t tell that from a still image.”

Some participants were able to imagine how they might like to have the video displayed, as well as how they might like to be able to manipulate the video. Participant #2, for example, stated that,” if I could pop through it…look at a piece of it…20 to 30 seconds and then fast forward…that would be great.” Participant #3 vocalized how the video might be made easy to scan through. “If the video was broken down, and there were representative frames and representative shots displayed. Then, if I could go into one of those to view the video…that would be very useful.” Similarly, participant #5 felt that for online lectures, it would be good “if there is a clear introduction or conclusion section that lasts a few seconds.”

Question #8 asked the participants whether or not they would like to be able to see written transcripts of the video at this point. Some felt that this would be of little or no use, however four of the participants thought it would be very useful for their needs.

Participant #2, who is a journalism professor, said that, “if I was showing a class a commercial, and it’s my writing class…if I could pop out the actual script…that would be very useful.” Participant #3 thought that transcripts would be very useful for an interesting reason. This person stated that in the making of documentaries, the transcript is “often used extensively in planning.” When other footage is brought in to their own work, often a transcript has to be made. “If there was already one available that we could download, that would save a lot of time.” Participant #7, felt that seeing a transcript would help alleviate some of the problems that may exist with audio quality. “I don’t know if you ever try to sometimes hear some of these things but the quality does degrade a little bit. So if you had a written transcript flowing by while I was watching the video…wow…then I am locked in! It’s all perfect at that point.” This person also said that having a transcript would be great because, “I could search!” Participant #9, who was involved in producing video, felt that being able to scan through a transcript would help in making a decision on whether or not to download a particular video.

Question 9 dealt with the issue of whether or not the participants would like to go through multiple phases of a search. They were asked if they would like to be able to separate into a subset videos that they felt might be of use, and then be able to browse this subset. All felt that this would be a valuable process. Participant #4 said, “Definitely. What I like is the idea, you know, you don’t want to spend a lot of time wondering if this is the right one. It would be good to say, ‘Yes. I want to look at this one, and this one, and this one,’ and then sort later.”

Question 10 aimed at seeing whether, once a retrieval set was returned, the participants would like to be able to use one item in the retrieved set to use as an example to search for similar videos. Most felt that this would be a useful activity; however, participant #9 felt that this process would not help in finding a specific video. Participant #7 thought that it could be useful, but expressed doubts as to whether a system would be able to perform this action successfully. “That’s an interesting thought. Querying by example seems powerful. I might want to try to do that. I just don’t know if I have a lot of confidence in the technique. It is hard in a multidimensional kind of data type like video, for the system to really know what about the video I want more similar parts of.”

Summary

These findings clearly indicate that there is a large need for video in many fields. These interviews only represent a fraction of the potential individuals that could have been interviewed for this study. It would be easy for more user studies like this one to be done, and this research is enthusiastically encouraged! While this is a very small sample of individuals, the fact that there were usage similarities within this small study suggests that it may be possible to establish concise user profiles. This would enable targeted video databases to be created for particular fields or industries.

Another finding that should be emphasized is that the participants in this study wanted to express themselves using words. Many felt that they had a good idea of the specific video that they wanted and that the best way to express that to a video retrieval system is

through words. Furthermore, the participants often wanted to specify their needs to a system using terms that were specific to their particular field or domain.

Once a video set is retrieved, these individuals seemed to want to see a lot of information to give them an idea of whether the retrieved videos were indeed relevant. While some individuals were more interested in knowing specific metadata attributes about the videos, most felt that it was very important to be able to view the actual videos. Videos are not text. People are interested in seeing the quality of the video as well as hearing the quality of the audio. This is particularly true for the individuals involved in video production or creative video work.

Conclusion

Establishing user needs is a vital part of a larger system development process and is key in what Preece (1999) refers to as the requirements gathering phase of system development. “Requirements gathering or analysis is the process of finding out what a client (or customer) requires from a software system. One of the main purposes of requirements gathering is to classify the clients’ needs and to identify infeasible requirements, omission, ambiguities and vagueness” (Preece, 1999). Hix and Hartson (1993) state that the process of needs analysis “establishes that a new system is in fact needed, based on goals of the organization and demands of the marketplace, and it determines the basic goals, purpose, and features desired for the application system. Features are characteristics and capabilities of the system as they appear to the users. The result is an external view of what a user will be able to do with the application system.”

(p. 117-118) With this in mind, I believe that, while the focused research currently being done by the video retrieval community is not intrinsically negative, it must be balanced by needs analysis in order for it to be useful. Only then, can realistic scenarios be built, which can later be used to develop usability studies that test the specific design being developed.

The problem of the lack of user needs consideration, it must be stated, is not only an aspect of video retrieval research. In fact, image retrieval research has suffered from the same lack of understanding of user needs and behaviors. Basically, most non-textual retrieval research can be seen as suffering from this problem. Eakins and Graham (1999) state that “Remarkably little has yet been published on the way such users search for and use images, though attempts are being made to categorize users' behaviour in the hope that this will enable their needs to be better met in the future.” Furthermore, “What kinds of query are users likely to put to an image database? To answer this question in depth requires a detailed knowledge of user needs - why users seek images, what use they make of them, and how they judge the utility of the images they retrieve…Not enough research has yet been reported to answer these questions with any certainty” (Eakins and Graham, 1999). It was the goal of the current study to address this problem.

The questions this research attempted to address are important ones. Having a good idea of why and how people want to use video retrieval systems is and will continue to be a huge part of developing successful systems. User-centered design is well established as a vital part of system design. This underlying idea of designing systems to meet the needs

of users is at the heart of many of the research projects currently being undertaken in the field of video retrieval. However, too many are based on assumptions about what users actually need or want rather than empirical data on users' information needs. More research needs to be done which goes directly to the user and asks the questions that were asked in this research: Why do you need video? What criteria do you think would be useful in specifying to a system how to find the type of video you need? What kind of information would help you decide, given a retrieval set, which videos would be the most relevant to your needs? When these questions are answered, a clearer picture can be made of what constitutes a successful system.

This study, though conducted with only a small number of potential users, takes a first step in providing information that designers can use in developing video retrieval systems. So, what does this suggest about designing a video retrieval search interface? Well, first of all, it suggests that designing a system for a broad, varied audience will be extremely difficult, especially when one considers attempting to index video without any idea of what aspects potential users will find important. Furthermore, the system must be flexible enough to allow for specific searches as well as users that would like to browse. As more needs analysis studies are done, it will become clearer to designers what users want and expect from a video retrieval system.

Bibliography

Belkin, N. J. (1984). Cognitive models and information transfer. Social Science Information Studies, 4, 111-130.

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982a). ASK for information retrieval: Part I. Background and theory. Journal of Documentation, 38, 61-71.

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982a). ASK for information retrieval: Part II. Results of a design study. Journal of Documentation, 38, 145-164.

Bias, R. & Mayhew D. (Eds.). (1994). Cost-Justifying Usability. San Francisco, CA: Morgan Kaufman Publishers.

Breton, E. J. (1981). Why engineers don't use databases. Bulletin of the American Society for Information Science, 7, 20-23.

Chang, Shih-Fu (Moderator). (1999). Multimedia access and retrieval: the state of the art and future directions. Proceedings of ACM Multimedia, 1, 443-445.

Chang, S., Chen, W., Meng, H. J., Sundaram, H., & Zhong, D. (1997). VideoQ: an automated content based video search system using visual cues. Proceedings of ACM International Conference on Multimedia. 313-124.

Christel, M. G., Smith, M. A., Taylor, C. R., & Winkler, D. B. (1998). Evolving video skims into useful multimedia abstractions. Proceedings of CHI '98, Conf. Human Factors in Computing System 171-178.

Cronin, B. (1981). Assessing user needs. Aslib Proceedings, 33, 37-47.

Dervin, B. (1983a). Information as a user construct: the relevance of perceived information needs to synthesis and interpretation. In R. Rice & W. Paisley, (Eds.), Knowledge Structures and Use: Implications for Synthesis and Interpretation (pp. 153-184). Philadelphia, PA: Temple University Press.

Dervin, B. (1983b). An overview of sense-making: concepts, methods, and results to date. Paper presented at International Communication Association Annual Meeting, Dallas, TX.

Dervin, B., & Nilan, M. (1986). Information needs and uses. Annual Review of Information Science and Technology (ARIST), 21, 149-160.

Ding, W., Marchionini, G. & Tse, T. (1997). Previewing video data: browsing key frames at high rates using a video slide show interface. Proceedings of the International Symposium on Research, Development & Practice in Digital Libraries, 151-158.

Dragon, R. Editorial review of Usability Engineering by Jakob Nielsen. [On-line], Available: http://www.amazon.com/exec/obidos/ASIN/0125184069/102-5706163-4403356

Durrance, J. C. (1984). Armed for Action – Library Response to Citizen Information Needs. New York, NY: Neal-Schuman Publishers, Inc.

Eakins, J. P. & Graham, M. E. (1999) Content-based image retrieval, a report to the JISC technology application programme. Technical report, Institute for Image Data Research, University of Northumbria at Newcastle, UK. Available: http://www.unn.ac.uk/iidr/report.html

Flicker, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., & Yanker, P. (1995) Query by image and video content: The QBIC system, IEEE Computer Magazine, 28, 23-32.

Ford, N. (1980). Relating “information needs” to learner characteristics in higher education. Journal of Documentation, 36, 99-114.

Garvey, W. D., Tomita, K., Woolf, P. (1979). The dynamic scientific-information user. Communication: The Essence of Science. (pp. 256-279). Elmsford, NY: Pergamon Press.

Geisler, G. & Marchionini, G. (2000). The Open Video Project: a research-oriented digital video repository [short paper]. In Digital Libraries '00: The Fifth ACM Conference on Digital Libraries (June 2-7 2000, San Antonio, TX), 258-259.

Geisler, G., Marchionini, G., Nelson, M., Spinks, R., & Yang, M. (2001). Interface concepts for the Open Video Project. University of North Carolina at Chapel Hill. Manuscript submitted for publication.

Hackos, J. T. & Redish, J. C. (1998). User and Task Analysis for Interface Design. Somerset, NJ: John Wiley & Sons, Inc.

Hamrapur, A., Gupta, A., Horowitz, B., Shu, C. F., Fuller, C., Bach, J., Gorkani, M., & Jain, R. (1997). Virage video engine. SPIE Proceedings on Storage and Retrieval for Image and Video Databases, 188-197.

Hansen, W. J. (1971). User engineering principles for interactive systems. AFIPS Fall Joint Computer Conference, (pp. 523-532). New Jersey: AFIPS Press.

Haskins, R. (1993) The Shark Continuous-Media file server. Proc. IEEE COMPCON '93

Hix, D. & Hartson H. R. (1993) Developing User Interfaces: Ensuring Usability Through Product and Process. Somerset, NJ: John Wiley & Sons, Inc.

Jain, R. (1997). Visual information management. Communications of the ACM, 40, 29-30.

Jain, R. & Hampapur, A. (1994). Metadata in video databases. SIGMOD Record (ACM Special Interest Group on Management of Data), 23, 27-33.

Jarvelin, K., & Repo, A. J. (1982). Knowledge work augmentation and human information seeking. Journal of Information Science, 5, 79-86.

Keeton, K., & Katz, R. (1993). The evaluation of video layout strategies on a high-bandwidth file server. Fourth Int'l Workshop on Operating Systems and Network Support for Digital Audio and Video

Krikelas, J. (1983). Information seeking behavior patterns and concepts. Drexel Library Quarterly, 19, 5-20.

Lienhart, R., Pfeiffer, S., & Effelsberg, W. (1997). Video abstracting. Communications of ACM, December '97 55-62.

Lowry, G. R., (1979). Information Use and Transfer Studies: An Appraisal. 37p.

MacMullin, S. E., & Taylor, R. S. (1984) Problem dimensions and information traits. The Information Society, 3, 91-111.

Maron, M. E. (1985) Probabilistic retrieval models. In B. Dervin & M. Voigt (Eds.) Progress in Communication Sciences. Norwood, NJ: Ablex Publishing.

Mayhew, D. (1999). The Usability Engineering Lifecycle. San Francisco, CA: Morgan Kaufman Publishers.

Mick, C., Lindsey, G. N., & Callahan, D. (1980). Toward usable user studies. Journal of the American Society for Information Science, 31 (5), 347-356.

Nielsen, J. (1994). Usability Engineering. San Francisco, CA: Morgan Kaufman Publishers.

Nielsen, J. (1998). [Comment about book], [On-line] Available: http://www.amazon.com/exec/obidos/ASIN/0125184069/102-5706163-4403356

Nielsen, J. (2000). Designing Web Usability. Indianapolis, IN: New Riders Publishing.

Norman, D. A. & Draper, S. W. (Eds.). (1986). User Centered System Design. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ofori-Dwumfuo, G. O. (1982). Reference retrieval without user query formulation. Journal of Information Science , 4, 105-110.

Paisley, W. (1980) Information needs and uses. In B. Dervin & M. Voigt (Eds.) Progress in Communication Sciences. 2, Norwood, NJ: Ablex Publishing.

Pearrow, M. (2000). Web Site Usability Handbook. Hingham, MA: Charles River Media.

Pentland, A., Picard, R. W., & Sclaroff, S. (1996). Photobook: content-based manipulation of image databases. International Journal of Computer Vision, 18, 233-254.

Pew Internet & American Life. (2001). More online, doing more [On-line]. Available: http://www.pewinternet.org/reports/toc.asp?Report=30

Ponceleon, D., Srinivasan, S., Amir, A., Petkovic, D., & Diklic, D. (1998). Key to effective video retrieval: effective cataloging and browsing. ACM Multimedia, '98, 99-107.

Preece, J. (1999) Human-Computer Interaction. Harlow, England: Addison-Wesley.

Rangan, P. V., Vin, H. M., & Ramanathan, S. (1992). Designing an on-demand multimedia service. IEEE Communications Magazine, 30 (7) 56-64.

Robertson, S. E. (1977). Theories and Models in Information Retrieval. Journal of Documentation, 33, 126-148.

Rosenfeld, L. & Morville, P. (1998). Information Architecture for the World Wide Web. Sebastopol, CA: O'Reilly & Associates.

Rowe, L., Boreczky, J. S., Eads, C. A. (1994). Indexes for user access to large video databases. Storage and Retrieval for Image and Video Databases II, IS&T/SPIE

Shneiderman, B. (1988). Designing the User Interface: Strategies for Effective Human-Computer Interaction. Reading, MA: Addison-Wesley Publishing Company.

Slaughter, L., Marchionini, G., & Geisler, G. (2000). Open Video: a framework for a test collection. Journal of Network and Computer Applications, 23 (3), 219-245.

Smith, J. R., Chang, S. F. (1996) VisualSEEK: A fully automated content-based image query system. ACM Multimedia Conference, 87-98.

Spool, J. (1999). Web Site Usability: A Designer's Guide. San Francisco, CA: Morgan Kaufman Publishers.

Tobagi, F. A., & Pang, J. (1993). StarWorks *TM – a video applications server Proc. IEEE COMPCON ’93

Vermuelen, C. H. (1981) What about the Non-Users of the Public Library? South African Journal for Librarianship and Information Science, 49, 10-14.

White, H. (1980). Library Effectiveness – The Elusive Target. American Libraries, 11, 683.

Williamson, N. J. (1981). Viewdata Systems: Designing a Database for Effective User Access. Canadian Journal of Information Science, 6, 1-14.

Wooster, H. (1979). The Users Last. Bulletin of the American Society for Information Science, 5, 19-21.

Appendix A: Interview Questions

1. What was your most recent experience with searching for video? What kind of video did you need and why? Was it a positive experience? Did you find what you were looking for?

2. In a few sentences, describe a realistic situation where you would possibly need to search for a video or video segment online? What kind of video would you need and why?

3. How important would specifying the following be in searching for your video?

Title: the name given to the video.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Creator: the person or organization that created the video.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Subject or keywords: terms that describe the video.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Contributor: the person or organization that contributed the video to the collection.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Creation date: the date the video was created.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Format: the digitized format of the video (example: MPEG, Real Video, Quicktime, etc.)
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Language: the language spoken in the video.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Rights Information: Copyright information and any restrictions of usage.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

File Size: the size of the video file.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Length: the length of the video file in minutes and seconds.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Amount of motion: the amount of motion contained within a video on a scale.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Color: whether or not the video is in color.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Sound: whether or not the video has sound.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

People: whether or not the video contains people.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Are there any other criteria you would find useful in specifying the kind of video you need? Using the same scale as above, how useful would it be?

For the video you are searching for, how useful would it be to be able to search using a technique other than either typing or clicking on words? For example:

Drawing your query: describing what you are looking for by drawing it.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Specifying colors: telling the system colors you require.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Specifying motion: somehow describing the motion you need in the video.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Are there any other non-textual techniques you can imagine would be useful in searching for the video you need? Using the same scale as above, how useful would it be?

4. For the next set of questions, imagine that you have input some kind of search criteria into the system and have received back a set of videos and corresponding information. How important would knowing the following be in determining whether a video is right for you:

Title: the name given to the video.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Creator: the person or organization that created the video.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Subject or keywords: terms that describe the video.

___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Contributor: the person or organization that contributed the video to the collection.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Creation date: the date the video was created.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Format: the digitized format of the video (example: MPEG, Real Video, Quicktime, etc.)
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Language: the language spoken in the video.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Rights Information: Copyright information and any restrictions of usage.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

File Size: the size of the video file.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Length: the length of the video file.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Amount of motion: the amount of motion contained within a video on a scale.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Color: whether or not the video is in color.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

Sound: whether or not the video has sound.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

People: whether or not the video contains people.
___ Very useful ___ Somewhat useful ___ Not very useful ___ No use at all

5. Regarding the above, would you rather retrieve video segments (smaller pieces of a video) and corresponding information or the entire video and corresponding information? If so, how would you like the video broken up? Using what criteria?

6. Would you want to be able to see still images of the video at this point? If so, how do you think you would like the images displayed?

7. Would you want to be able to see video clips at this point? If so, how do you think you would like the video displayed?

8. Would you want to be able to see written transcripts of the video if they were available at this point?

9. Would you want to be able to separate the videos you think you are interested in from the videos you think you are not interested in at this point?

10. Would you like to be able to use specific videos as examples in order to “search for similar videos?”